Contention-based Limited Deflection Routing in OBS Networks

SuKyoung Lee*, Kotikalapudi Sriram†, HyunSook Kim‡ and JooSeok Song‡
*Sejong University, Seoul, Korea †NIST, Gaithersburg, USA ‡Yonsei University, Seoul, Korea

***Abstract*— One of the critical design issues in Optical Burst Switching (OBS) networks is finding ways to minimize burst dropping resulting from resource contention. The existing variants of the basic deflection routing schemes all lack the ability to determine the alternate route based on clear performance objectives. In this paper, we present Contention-Based Limited Deflection Routing (CLDR) scheme, which sequentially performs the following: (1) based on certain performance criteria, dynamically determines if the burst should be deflection routed or retransmitted from source, (2) if the decision is to deflection route, then the same is done using a path that is based on minimization of a performance measure that combines distance and blocking due to contention. Through analytical and simulation modeling, a number of useful insights into the OBS network protocols and performance are provided.**

I. Introduction

There is growing consideration for OBS for economic implementation of future all optical IP networks [1]. One of the challenging issues in the implementation of burst switching is the resolution of contentions that results from multiple incoming bursts that are directed to the same output port. Deflection routing is one of the promising ways for burst contention resolution. Some recent work on deflection routing is reported in [2]-[5]. The authors of [2] investigate the performance of deflection routing in prioritized Just Enough Time (JET)-based OBS networks. In [3] and [4], it is demonstrated via simulation tests that the blocking probability improves when deflection routing is used as a means for contention resolution. The authors of [5] describe how deflection routing can be used in conjunction with the self-routing address scheme. However, they do not address the issue of how routing to an alternate path should be done, given that some constraints may apply to the selection of an alternate path.

In this paper, we propose and analyze a novel Contention-based Limited Deflection Routing (CLDR) protocol, which mitigates and resolves contention with significantly better performance as compared to techniques currently known in the literature. While several variants of the basic deflection routing scheme have been proposed before [2]-[5], they all lacked the ability to determine the alternate route based on clear performance objectives. In this paper, we present an on-demand deflection routing scheme, which sequentially performs the following: (1) based on certain performance criteria, dynamically determines if the burst should be deflection routed or retransmitted from source, (2) if the decision is to deflection route, then the same is done using a path that is based minimization of a performance measure that combines distance and blocking due to contention. The proposed CLDR scheme prevents injudicious deflection routing. Our simulation results show that the scheme proposed here has much superior performance both in terms of burst loss probability and increased network throughput. In this paper, we have also proposed that the network nodes should periodically re-compute and store optical paths, with the aim of staying optimal in the face of changing node and link congestion measures. This allows for deflection routed bursts to traverse the alternate optical paths that are not necessarily shortest path but are optimized for best performance (i.e., blocking and delay). This technique calls for monitoring the link and node congestion and updating the same in a periodic manner so that the path computation can be as optimal as possible (albeit with some minor lag in the updates).

Further, we have presented here an analytical model for computation of burst loss ratio due to contention on congested links in the network. Typically, the traffic originating from the edge nodes of the network would be correlated and such correlations would have a significant impact on the burst contentions at the edge as well on internal links in the network. Our analytical model accounts for these correlations (including various parameters that help quantify the correlations) in the prediction of burst loss ratio or probability. The analytical model results are compared with simulation results. Additionally, the analytical modeling results also used to create some relevant inputs in the design of the simulation experiments for studying CLDR and comparing it with other known schemes.

II. Contention-based Limited Deflection Routing Algorithm

Several reservation protocols have been proposed recently for implementing optical burst switching with different wavelength and timeslot reservation schemes. Two examples of such protocols are: (1) offset-based scheme including JET [6] and Just-In Time (JIT) [7] and (2) FDL-based scheme [8]. The proposed CLDR mechanism can be applied to both styles of burst scheduling/reservation stated above, i.e., with offset time or with FDL. This is so because the CLDR is not a mechanism for how to reserve the wavelength and timeslot but it is about a criterion for deciding on doing deflection routing and further about a methodology for selection of the alternative path.

A. Performance Measurement Updates

Fig. 1 shows a flowchart describing the operation for performance measurement updates in the CLDR scheme. There exists a management database referred to in Fig. 1 as the Deflection Routing Information Base (DRIB) at the OBS edge node. The DRIB stores the management information for the optical burst layer together with the traditional DWDM transport and IP layers of network. An updated measurement about burst contentions is needed at all the nodes in the network for the CLDR algorithm to perform well. The flow chart of Fig. 1 illustrates the mechanism for signaling contention occurrences and updating the burst contention measurement. Each ingress node receives updates about the burst congestion status along the primary and alternates routes for the bursts that have originated from it. These updates come in the form of two kinds of $NACK$ messages, $NACK\_C$ and $NACK\_D$: the $NACK\_C$ message is sent with an incremented value when

0-7803-7975-6/03/$17.00 (C) 2003

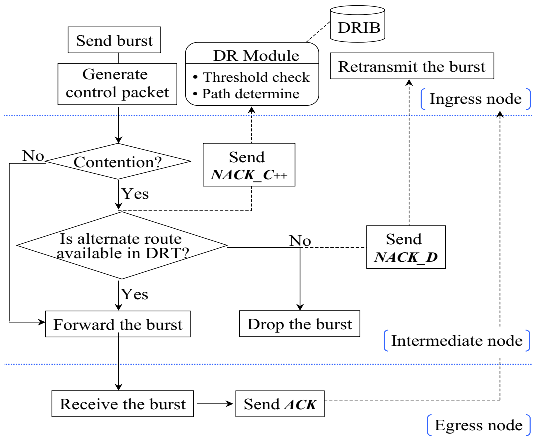

Fig. 1. A flowchart describing burst contention notification and measurement in the CLDR algorithm

contention occurs on the primary path due to the lack of a timeslot in a wavelength while $NACK\_D$ is also sent when there is no available alternate route in the Deflection Routing Table (DRT). These messages help update the DRIB at the ingress nodes of each burst-mode connection.

B. Computation of Alternate Routes

In an OBS network, a local routing decision for the alternate route made at a node may result in a degraded global network performance in the long run. However, this is mitigated in the proposed CLDR algorithm by performing periodical global re-optimization of alternate routes based on updates received from other nodes regarding their most recent contention status (Fig. 1).

We now formulate the deflection routing problem by means of the following components: the network topology, node configuration, a set of attributes pertaining to node and link resources, and constraints pertaining to resource limits. The demands that are to be routed through alternate paths in the network are described by a set of attributes as well. Then, the problem is to find an optimal alternate path minimizing a cost function, which explicitly accounts for the contention rate as well as the burst hop distance. The aforementioned deflection routing problem can be formulated as follows: Consider a physical network represented by a graph $G(N, L)$, where N is the set of nodes and L the set of links (*i.e.*, fibers) connecting the nodes. It is assumed that each link between nodes i and j, has W_{ij} wavelengths with the same capacity of C. At each node n, $(n = 1, \ldots, N)$, the number of transmitters and receivers are defined as $P_n^{(t)}$ and $P_n^{(r)}$, respectively. If a node n has the number, P_n of ports, clearly, at most $\sum_n P_n$ wavelengths are needed to realize any possible virtual topology.

Let Λ be the set of traffic demands belonging to the loss-sensitive service class between a pair of edge nodes, where $\lambda_{ij}^{sd} \in \Lambda$ represents the arrival rate of bursts from source s to destination d that flows over a virtual link between node i and node j. Further, let $\lambda_{s_k d_k}$ denote the average flow of bursts associated with the k^{th} traffic demand requesting service. Let $D = \{D_{ij}\}$ be the distance matrix from node i to node j $(i \neq j)$. As the cost of contention from node i to node j $(i \neq j)$, let b_{ij} denote the burst blocking rate, which is collected periodically from the network.

In the deflection routing problem formulation, the variable, x_{ij} is defined as

$$x_{ij} = \begin{cases} 1 & \text{if alternate route includes a link } (i, j) \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

where $i, j = \{1, 2, \ldots, N\}$ and $i \neq j$. This decision variable x_{ij} pertains to the specific k^{th} traffic demand at hand which is characterized by the $\lambda_{s_k d_k}$ average flow of bursts. Here, for the purpose of routing decisions, we are treating each burst-oriented Variable Bit-Rate (VBR) connection request as a Constant Bit-Rate (CBR) connection with an effective bandwidth of $\lambda_{s_k d_k}$.

The constraint conditions are defined as follows. The number of lightpaths originating from and terminating at a node is not more than the node's out-degree and in-degree, respectively:

$$\sum_{\forall j \in N} x_{ij} \leq P_i^{(t)}, \sum_{\forall i \in N} x_{ij} \leq P_j^{(r)} \qquad (2)$$

There are some constraints related to the traffic flow on virtual topology for all i and j. First, since we are setting up an alternate path for the optical bursts coming from a specific traffic flow the bursts of the demand $\lambda_{s_k d_k}$ are not segmented at any congested node in the network. Thus, we can state that the traffic demand $\lambda_{s_k d_k}$ is routed from node i to j on a single deflected path:

$$\lambda_{ij}^{s_k d_k} \in \{0, \lambda_{s_k d_k}\} \qquad \forall i, j \in N \qquad (3)$$

Second, the total flow on the simplex link from node i to node j is expressed as the superposition of the existing traffic (*i.e.*, bursts) and the new burst associated with the link.

$$\lambda_{ij} = x_{ij} \sum_{s,d} \lambda_{ij}^{sd} + x_{ij} \lambda_{s_k d_k} \qquad \forall i, j \in N \qquad (4)$$

For flow conservation at each node, the third constraint becomes

$$\sum_j x_{ij} - \sum_j x_{ji} = \begin{cases} 1, & i = s_k \\ -1, & i = d_k \\ 0, & \text{otherwise} \end{cases} \quad \forall s_k, d_k, i \in N \quad (5)$$

With regard to traffic, the final constraint assures that traffic flowing through a link can not exceed the total link capacity:

$$\lambda_{ij} \leq W_{ij} C \qquad \forall i, j \in N. \qquad (6)$$

Then, for the above constraints and the k^{th} traffic flow of loss-sensitive service class, we can design the objective function to find an alternate path from the congested node to the destination. This objective function is a weighted sum of the end-to-end burst blocking rate and the distance for the route, and is stated as follows:

$$Minimize\ g_d \sum_{i,j} x_{ij} D_{ij} + g_b [\log_{10} [1 - \prod\nolimits_{i,j} (1 - x_{ij} b_{ij})] \quad (7)$$

where g_d, and g_b denote the weights for delay and blocking, respectively. To decrease the computational complexity, we can express the above objective function, Eq. 7 as

$$Minimize \qquad g_d \sum_{i,j} x_{ij} D_{ij} + g_b \sum_{i,j} x_{ij} \log_{10} b_{ij}, \qquad (8)$$

As for the burst contention rate, b_{ij}, the real data can be used that has been collected into the DRIB. The alternate route would be set up according to the values of the x_{ij} determined from the above integer linear programming formulation. The objective function in Eq. 8 is more of a practical value than one involving distance alone. It includes Quality of Service (QoS) requirements regarding loss as well as distance.

C. Limited Deflection Routing Rules for CLDR

Our CLDR algorithm consists of (1) the optimal alternate routing methodology for Contention-based Deflection Routing that was described in the preceding section and (2) the rules of Limited Deflection Routing (LDR) that we describe in this section. The authors of [4] proposed an LDR that is based on a simple sender check function before deflection routing at the source node only. We propose an enhanced LDR with a comprehensive threshold check function, described in the following steps: (1) Source node sends out a control packet; (2) Intermediate nodes process the control packet and attempt to reserve the channel in anticipation of the burst that would follow; (3) Source node sends out the burst after offset time; (4) If there is no available egress channel for the burst at a node, at first it is checked whether the current node is sender or not. If the current node is the sender, then deflection routing is not done. Instead, after some wait time, the sender retransmits a burst control packet and subsequently the burst is retransmitted. If the current node is an intermediate node, then go to the next step (5); (5) The current intermediate node computes a performance measure and does the threshold check on that performance measure. Accordingly, it decides whether to deflection route or drop and notify sender to retransmit.

The enhanced LDR described above reduces unnecessary deflection routing at the intermediate nodes as well as at the sender, and prevents contentions which are caused by inefficient deflection routing.

Now, we formulate some threshold check functions to assist in deciding whether dropping or deflection routing should be done. Let N be the set of nodes in core network. Further, let N_c and N_d be the set of nodes that have been passed from source to the current node (i.e., the congested node in consideration) and the set of nodes that would be passed from the current node to destination, respectively. $x_{i,i+1}$ is a binary variable associated with link $(i, i+1)$ between a node i and the next node $i+1$.

We first define a component of threshold check function, which is based on hop counts:

$$C_h = \sum_{\forall i,i+1 \in N_c} x_{i,i+1} - \sum_{\forall j,j+1 \in N_d} x_{j,j+1} \tag{9}$$

Let b^* denote a tolerable end-to-end blocking rate for a route. We now define another components of the threshold check function to satisfy b^*:

$$C_b = \log_{10} b^* - \log_{10}[1 - \prod_{i=1}^{d-1}(1 - b_{i,i+1})] \; \forall i, i+1 \in N_d \tag{10}$$

where d is the destination node and $b_{i,i+1}$ denotes contention blocking probability between node i and node $i+1$.

Now we define a generalized threshold check function to include the path hop-count (or alternatively, the end-to-end distance) as well as the end-to-end burst blocking probability. The two performance measures in this threshold check are given different relative weights to emphasize one or the other, as desired. With a relatively large value M and burst blocking decision parameters b_2^* and b_1^* ($b_2^* > b_1^*$), we introduce two decision variables:

$$Q_h = \begin{cases} 1 & \text{if } C_h \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

and

$$Q_b = \begin{cases} 1 & \text{if } b_1^* \leq C_b \leq b_2^* \\ M & \text{if } C_b < b_1^* \\ -M & \text{if } C_b > b_2^* \end{cases} \tag{12}$$

Using the above two variables, we can now express a composite decision variable as

$$Q_t = w_h Q_h + Q_b \tag{13}$$

where $w_h \ll M$ is weight for emphasizing/de-emphasizing the hop count relative to the burst loss ratio. Then, a combined threshold check function can be stated as

$$C_t = \begin{cases} 1 & \text{if } Q_t \geq w_h + 1 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

According to the threshold check function 14, the proposed CLDR is capable of operating in the most suitable and efficient way under different traffic and topological scenarios.

III. Queueing Model for Burst Loss Probability

Numerical results showing performance of CLDR and its comparison to other schemes in OBS networks have been generated using both analytical and simulation modeling.

A. Source Traffic Model

We assume that a threshold-based burst assembly scheme is applied at the edge node, where once the length of a burst being created reaches a threshold value, L(Mbits), the burst is generated and placed in a burst queue. The incoming traffic on each wavelength of an incoming fiber is an aggregation of many individual sources of burst traffic. The bursts from each source are assumed to behave in an on-off manner. Many bursts arrive with exponential random inter-arrival times or at fixed intervals during the on period, and the off period is typically much longer than the on period. As an example on average 12 bursts may arrive during the on period over 120 ms average duration, and the off period may have 880 ms average duration. We varied these parameters widely in our analytical and simulation results so as to capture the sensitivity to these parameters.

B. Queueing Model

We closely follow the analytical model presented in [9] for queueing analysis involving multiplexing of many on-off sources, with suitable modifications for the statistical burst multiplexing problem at hand. To describe the analytical model, let us define the following parameters:

- C: link capacity (Gbps)
- α^{-1}, β^{-1}: average on and off periods, respectively (ms)
- λ: burst generation rate during on period
- n: number of sources simultaneously multiplexed on a link
- B: the burst queue size or the number of FDLs per output port (specified in total ms worth of buffering at link speed)
- i: system state in terms of number of sources simultaneously in on period($0 \leq i \leq n$)

- τ_i: effective time (ms) spent in system state i for burst delay to exceed buffer size Bms
- p_i: probability that system is in state i
- n_0: number of sources in on period simultaneously above which the system is considered to be in temporary overload

Following the rationale of [9], we compute the duration of time τ_i that the system needs to be in state i for burst delay to exceed B ms:

$$\tau_i = \frac{\frac{BC}{L\lambda}}{(i-n_0)(i-n_0+1)}. \tag{15}$$

Now, an approximate expression for the probability of burst loss due to buffer overflow P_L is given as follows:

$$P_L = \sum_{i=n_0+1}^{n} p_i \times \exp^{-i\alpha\tau_i} \tag{16}$$

where $p_i = \begin{pmatrix} n \\ i \end{pmatrix} (\frac{\beta}{\alpha+\beta})^i (\frac{\alpha}{\alpha+\beta})^{n-i}$ are system-state probabilities. The above equation for burst blocking is very useful in that, unlike many other approximations available in the literature, it captures the effects of numerous traffic/system parameters as well as the influence of temporal correlations in the superposition of on-off sources [9][10].

IV. Performance Results and Comparisons

A. Analytical Model Based Performance Results

In our numerical results, for example, we have used C = 6 Gbps (six wavelengths of 1 Gbps each) and L = 1 Mb. Fig. 2 shows the sensitivity of the burst loss ratio to the average on and off periods. A comparison is shown for the cases of (α^{-1} = 120 ms, β^{-1} = 880 ms) vs. (α^{-1} = 12 ms, β^{-1} = 88 ms); both cases have an average activity factor of $a = 0.12$. What we see is that while the activity factor is constant, the burst loss ratio is higher for the case of higher average on period.

What is very interesting in Fig. 3 is that for the no FDL case, the burst loss behavior is entirely the opposite of what was stated in the previous paragraph involving Fig. 2. Burst loss ratio in Fig. 3 is in fact lower at low to moderate loads for the case of the larger on period. Only when the buffer size increases, do we see that the burst loss ratios go higher for the larger on period in the low to moderate load region (the graphs for different buffer sizes were omitted due to space limit). This again can be explained by a combination of these observations: (1) at a given percentage link load level, the number of burst sources multiplexed are 2.5 times more for the case of $a = 0.12$ as compared to that for $a = 0.3$, and (2) the temporal correlations in the superposed burst traffic is not manifest at low to moderate loads and smaller FDL sizes, while it is quite influentially manifest at high loads with larger queues or buffer sizes.

B. Simulation Environment

Simulation tests were done for the CLDR algorithm and Shortest Path-based Deflection Routing (SPDR) algorithm which has been generally used in other known works [3]-[5]. In our simulation, the JET method of offset-based reservation is used and the burst sources were individually simulated with the on-off model as explained in the Queueing Model section. The tests were carried out using a 14-node NSFNET topology as shown in Fig. 4. The transmission rate of each link is 6 Gbps, consisting of six wavelengths each operating at 1 Gbps. Over NSFNET topology, five source-destination node pairs were chosen randomly and optical bursts are generated from the source nodes. Just as an example, looking at Fig. 4, let us say that some bursts whose source and destination are CA1 and NJ, respectively, experience burst contention at UT node on UT-MI link on the primary path (CA1-UT-MI-NJ). In our simulation, let us say that the DRT lists the (UT-CO-MN-IL-PA-NJ) and (UT-CO-TX-MD-NJ) as alternate candidate paths. Of these, (UT-CO-MN-IL-PA-NJ) is the shortest alternate path from UT to NJ. However, the CLDR scheme can very well select (UT-CO-TX-MD-NJ) as the preferred alternate path if that happens to be the only one that meets the requirement on combined (distance and blocking) performance objective.

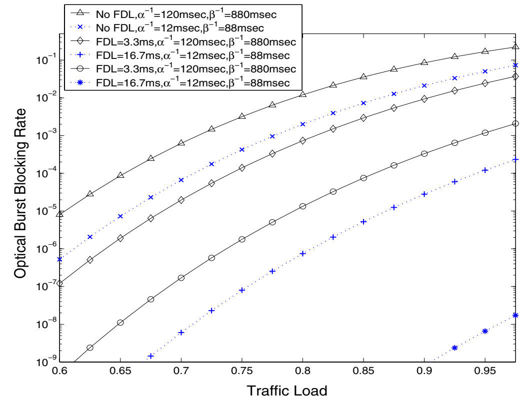

Fig. 2. Burst blocking rate for high and low on/off periods when $a = 0.12$

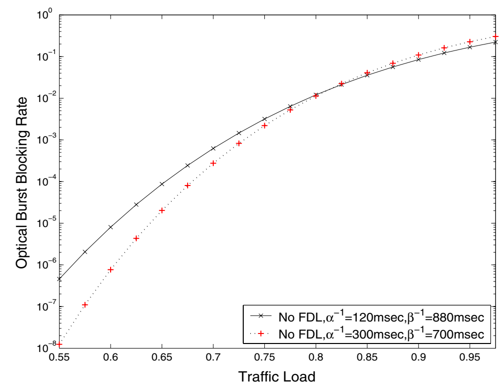

Fig. 3. Burst blocking rate sensitivity to on/off periods when there is no FDL

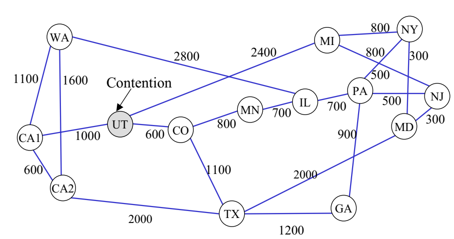

Fig. 4. Simulation network topology (in km)

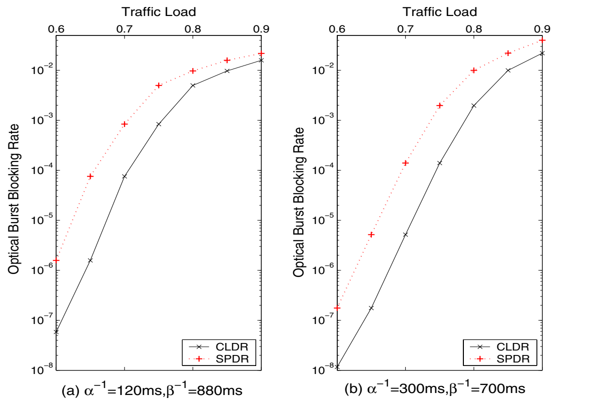

Fig. 5. Burst blocking rate for CLDR and SPDR without FDL when activity is 0.12 and 0.3

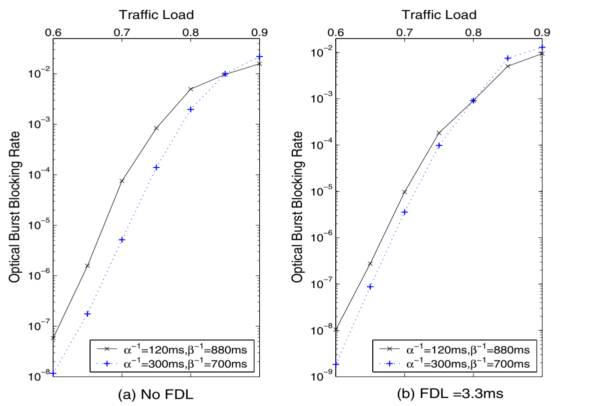

Fig. 6. Burst blocking rate for CLDR under two different cases of on/off periods and two cases of FDL sizes

C. *Performance of CLDR and Comparisons with Other Techniques*

The focus of our performance evaluation is on burst (or data) loss rate caused by contention. A burst would be dropped if both primary and deflection paths are blocked. The data loss rate for the entire network is found by calculating the average of the burst drops over all source-destination pairs.

Figs. 5(a) and (b) show simulation results comparing the burst blocking or loss rate for our CLDR method with the Shortest Path Deflection Routing (SPDR) method. For typical operating load values up to 0.75, the CLDR algorithm improves burst blocking by more than an order of magnitude as compared to SPDR in the test cases that we have studied through simulation runs.

Figs. 6 (a) and (b) show comparisons of the burst blocking ratios for the CLDR scheme under two different scenarios, namely, (α^{-1} = 120 ms, β^{-1} = 880 ms) and (α^{-1} = 300 ms, β^{-1} = 700 ms) for the cases of (a) no FDL and (b) 3.3 ms FDL. The reason we made this study of parameter sensitivity was to see if the burst loss results from simulations would be qualitatively in alignment with the analytical results of Fig. 3, and it turns out they are.

Now we turn to the second aspect of CLDR, namely, the enhanced limited deflection routing decision, based on a threshold check function, which is performed at an intermediate node experiencing burst contention. From Fig. 7, we see that the CLDR has better burst loss performance by about an order of magnitude as compared to the normal DR scheme, both under moderate and high link loads.

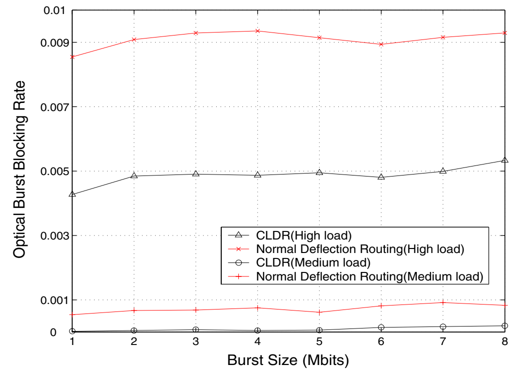

Fig. 7. Burst loss rate comparison for normal deflection routing vs. CLDR under moderate and high load conditions

V. Conclusion

In this paper, we have shown that, in OBS networks, when deflection routing is used as a means for burst contention resolution, it is important to design alternate routes in an optimized fashion based on a composite performance measure that considers path distances as well as the expected burst loss probability along that alternate route. An additional salient feature of CLDR is that the limited deflection routing decision is based on a threshold-check function and is dynamically made at any intermediate node experiencing burst contention. The proposed CLDR scheme was shown to perform significantly better than the SPDR scheme or other variants of the DR scheme that are known in the literature.

References

[1] K. Sriram, D.W. Griffith, S.K. Lee, and N.T. Golmie, "Optical Burst Switching: Benefits and Challenges," to appear, *International Workshop on Optical Burst Switching*, Dallas, USA, Oct. 2003.

[2] C-F. Hsu, T-L Liu and F-F Huang, "Performance Analysis of Deflection Routing in Optical Burst-Switched Networks", *IEEE Proceedings INFOCOM'02*, vol. 1, pp. 66-73, NewYork, USA, Jun. 2002.

[3] S. Kim, N. Kim, and M. Kang, "Contention Resolution for Optical Burst Switching Networks Using Alternative Routing", *IEEE Proceedings ICC'02*, vol. 5, pp. 2678-2681, NewYork, USA, Apr/May. 2002.

[4] X. Wang, H. Morikawa, and T. Aoyama, "Burst Optical Deflection Routing Protocol for Wavelength Routing WDM Networks", *Optical Networks Magazine*, Nov./Dec. 2002, pp. 12-19.

[5] C.Y. Li et al, "Deflection Routing in Slotted self-routing Networks with Arbitrary Topology", *IEEE Proceedings ICC'02*, vol. 5, pp. 2781-2785, NewYork, USA, Apr./May 2002.

[6] C. Qiao, and M. Yoo, "Optical Burst Switching (OBS) - A New Paradigm for an Optical Internet," *Journal of High Speed Networks*, vol. 8, no. 1, pp. 69-84, Jan. 1999.

[7] J. Y. Wei, J. L. Pastor, R. S. Ramamurthy, and Y. Tsai, "Just-In-Time Optical Burst Switching for Multi-wavelength Networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 18. no. 12, pp. 2019-2037, Dec. 2000.

[8] C. Gauger, K. Dolzer, and M. Scharf, "Reservation strategies for FDL buffers in OBS networks", *Report*, IND, University of Stuttgart, 2001

[9] K. Sriram, T.G. Lyons, and Y.T. Wang, "Anomalies due to Delay and Loss in AAL2 Packet Voice Systems: Performance Models and Methods of Mitigation," *IEEE JSAC*, vol. 17 , no. 1, pp. 4-17, Jan. 1999.

[10] K. Sriram and W. Whitt, "Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data," *IEEE JSAC*, vol. SAC-4, no. 6, pp. 833-846, Sep. 1986.